THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2021-11
CY Cergy Paris Université, France

Consumers' welfare and compensating variation: survey and mode choice application

Paolo Delle Site, André de Palma, Karim Kilani

March 2021

CY Cergy Paris Université

cnrs

Consumers' welfare and compensating variation: survey and mode choice application

Paolo Delle Site[a,*], André de Palma[b], Karim Kilani[c]

[a] *Niccolò Cusano University, Rome, Italy*
[b] *CY Cergy Paris Université, France*
[c] *LIRSA, CNAM, Paris, France*

Abstract

We study the welfare change from project and policies when consumers' behaviour is described with additive random utility models. We consider the random compensating variation mainstream approach and review the latest methodological developments. The expectation of the random compensating variation is used as a measure of the average welfare change. Without income effect, it is expressed by the monetized difference of the expectations of the maximum utilities with and without the changes in monetary costs or quality. This measure reduces for the multinomial logit model to the logsum formula. More generally, the expectation of the compensating variation can be expressed as a path-independent line integral. The rule-of-a-half is an approximation of this line integral. With income effect, the expectation of the compensating variation, both unconditional and conditional on the choices without and with the changes, is provided by one-dimensional integrals which can be computed numerically. In the conditional case, the average welfare change is attributed to those keeping and those changing alternative. The cumulative distribution function of the compensating variation allows the analysis of inequalities by extending the classical Lorenz curve and Gini coefficient. This analysis is perfomed distinctly for positive and for negative values of the compensating variation. Treatment of observed and unobserved heterogeneity is included. The survey of theoretical results is illustrated with a numerical example in the context of transportation mode choice, based on large-scale data collected in France.

Keywords: random utility model; compensating variation; Gini coefficient; Lorenz curve; rule-of-a-half.

JEL codes: D11, D30, D60, R41, R42, R48

*Corresponding author

Email addresses: paolo.dellesite@unicusano.it (Paolo Delle Site), andre.de-palma@cyu.fr (André de Palma), karkil2205@gmail.com (Karim Kilani)

1. Introduction

1.1. Motivation and contribution

The paper deals with the measurement of consumers' welfare. The analysis is aimed at the evaluation of the changes in welfare from projects and policies when consumers' behaviour is described with discrete choice additive random utility models (ARUMs). This is of relevance to many fields. In transportation, a classical problem is the evaluation of the benefits that accrue to users as a consequence of mode choice responses to the project or policy. The analysis is developed with this specific application aim.

Changes, induced by the project or policy, in both monetary cost and quality, e.g. travel time, of the alternatives are of interest. We use the terminology state *without* the change and state *with* the change. Alternative wordings are *before*, or *ex-ante*, and *after*, or *ex-post*, states. In transportation analysis, we make a series of comparisons of two scenarios, usually the do-nothing case versus the do-something case. Each comparison is referred to a given time frame. Therefore, no temporal sequencing is envisaged for the two states.

Classical microeconomics, which treats divisible goods with deterministic utility, proposes several measures of welfare change. These include consumer's surplus and Hicksian measures, i.e. the compensating and the equivalent variations. Measures of welfare change for models that represent the consumption of discrete, mutually exclusive, goods with random utility, are adaptations of the measures of classical microeconomics.

After the econometrics of the main ARUM, the logit, was set up in the Seventies (McFadden, 1974), researchers defined measures of welfare change. Detailed citations are in the next section. They started from the case without income effect, i.e. the case where income plays no role in the discrete choice. This case is consistent with the traditional approach of transportation demand modelling, which rules out income effect on the basis of the assumption that expenditure on transportation is a low fraction of income. Different approaches were envisaged. The case with income effect is more involved and was tackled only later. According to economic theory, transportation modes should be priced according to the full social costs. Higher prices than current practice make income effect of increasing relevance, because individual responses to price changes are, in principle, income-dependent.

Today, the mainstream, consolidated approach, without and with income effect, is based on the random compensating variation (McFadden, 1999). The expectation of the compensating variation is assumed as average welfare change for a population of individuals who are homogenous with respect to systematic utilities. Although equally theoretically valid, the random equivalent variation is neglected in the evaluation practice, apparently without justification.

Differently from previous surveys (Jara-Díaz and Farah, 1988; Jara-Díaz, 2007; Karlström, 2014), the paper reviews the latest developments in the mainstream approach. These pertain, in particular, to the analytical derivation of transition probabilities and expectations of the compensating variation conditional on the transitions. These are of relevance because they provide methods to estimate the shares of shifters and non shifters, and the attendant average benefits. In addition, the paper provides results concerning the distribution of the compensating variation. This is of relevance for a twofold reason. It provides the means to compute analytically the shares of winners (values of the compensating variation higher than zero) and losers (values of the compensating variation lower than zero). It makes inequality analysis possible by the construction of the Lorenz curve and the Gini coefficient. This construction is new to the best of our knowledge and is provided by the present paper. The case of a distinct inequality analysis for losses and for gains is dealt with. A second new theoretical contribution is the treatment of observed and unobserved heterogeneity. This is of relevance when the individuals are heterogenoeus and welfare analysis for the aggregate of the population is to be performed.

1.2. Literature overview

The exposition adopts a chronological criterion.

The first methodological contributions on welfare change measurement with discrete choice ARUMs related to the case without income effect. Different approaches were proposed. Williams (1977), McFadden (1981) and Anderson et al. (1992) used measures based on the representative consumer approach. Anderson et al. (1992) suggest the following interpretation. The behaviour of N consumers with different tastes is described by the choices made by a single individual who has a preference for variety. The representative consumer makes N trips. The frequency of choice of one alternative is provided by the probability associated with the ARUM. The representative consumer's indirect utility equals N times the sum of the individual income and the expectation of the maximum utility.

The measure of welfare change is the surplus variation of the representative consumer. The approach yields for the welfare change N times the monetized difference of the expectations of the maximum utilities in the state with change and in the state without change. For multinomial logit, the formula reduces to N times the monetized difference of logsums. Since it is based on the expectation of the maximum utility, the measure is a utilitarian welfare function which retains both positive and normative dimension. The representative consumer approach was applied to the case with income effect by Delle Site (2013), who derived for multinomial logit a representative consumer's indirect utility.

Small and Rosen (1981) proposed for ARUMs without income effect a measure based on the aggregate compensating variation derived from the variation of the expenditure function. They obtained a line integral of choice probabilities in systematic utilities between the two states without and with a policy change. This integral is monetized by dividing by the constant marginal utility of income.

The rule-of-a-half is an alternative approach used extensively by practitioners. It was first proposed based on an intuitive justification in a planning study in the UK in the Sixties (Tressider et al., 1968). It provides a simple formula for the assessment of users' benefits for any demand model in the case without income effect. The rule-of-a-half is also used to attribute conventionally the benefits to transportation alternatives, to non-shifting users and new users of each alternative, and to components of the generalized cost of travel. Williams (1976) showed that, for any demand model, the rule-of-a-half can be derived with appropriate assumptions and approximations from a surplus measure with respect to generalized cost. Jara-Díaz (1990) specialized the derivation for ARUMs considering price and quality changes, starting from the line integral obtained by Small and Rosen (1981). The accuracy of the rule-of-a-half approximation with multinomial logit was investigated numerically by Ma et al. (2015).

To estimate welfare change with income effect, McFadden (1999) introduced the definition of random compensating variation. It is the income adjustment that equates maximum utility in the state with change to maximum utility in the state without change. He proved that, without income effect, the expectation of the compensating variation reduces to the monetized difference of the expectations of the maximum utilities in the state with change and in the state without change. For multinomial logit, it boils down to the monetized difference of logsums. Also, he proved that the expectation of the compensating variation is equal to the line integral of Small and Rosen (1981). The combination of this result with the one by Jara-Díaz (1990) suggests that the rule-of-a-half can be derived analytically as an approximation of the mainstream measure.

The case with income effect was approached first by simulation, i.e. by drawing from the distribution of the random terms (McFadden, 1999; Herriges and Kling, 1999; Cherchi et al., 2004). Later on, researchers developed analytic approaches. Relatively simple formulas for ARUMs have been provided, on the basis of the distribution of the compensating variation, by Karlström and Morey (2004) and Karlström (2014), and by de Palma and Kilani (2011). Dagsvik and Karsltröm (2005) derived a more involved formula from Hicksian, i.e. compensated, choice probabilities. de Palma and Kilani (2011) derived their formula from Marshallian, i.e. observed, choice probabilities. Additionally, these last authors provided formulas for the transition probabilities and for the expectations of the compensating variation conditional on the transitions. Closed-form expressions of the transition probabilities for multinomial logit only had been provided in a former paper (de Palma and Kilani, 2005).

Zhao et al (2012) and Delle Site and Salucci (2013; 2015) relaxed the assumption that random terms are unchanged in the state without and in the state with change. Delle Site (2021) investigated the large sample properties of the estimator of the monetized difference of logsums.

Additional methodological contributions to welfare analysis for discrete choice models relate to non ARUMs, i.e. to models with unrestricted unobserved heterogeneity where a vector of random terms enter each alternative's utility function in an unspecified way (Bhattacharya, 2015 and 2018). Since the practice of transportation demand modelling is based mostly on ARUMs, these developments are not included in the following sections.

1.3. Organization of the paper

The remainder of the paper includes the following. Section 2 reports on the main assumptions and identities related to discrete choice ARUMs. Sections 3-6 deal with welfare change. The expressions of the expectation of the compensating variation are made available in Section 3. The expressions of the transition probabilities and of the conditional expectations of the compensating variation are provided in Section 4. In Section 5, the support and the cumulative distribution function of the compensating variation are given, and the application of the Lorenz curve and Gini coefficient for inequality analysis is discussed. Observed and unobserved heterogeneity is tackled in Section 6. New results are in the form of propositions. The theoretical insights are illustrated with a numerical example in Section 7. Conclusions are in Section 8.

Appendix A includes, for the convenience of researchers and practitioners, a summary of key findings from classical microeconomic theory of divisible goods. Appendix B presents the microeconomic foundation of the econometric specifications introduced in Section 2 and used in the other sections of the paper.

2. Main assumptions and identities

2.1. Additive random utility models

Consider an individual facing the choice among a finite number of J alternatives, $J \geq 2$. In ARUMs, each alternative j is associated with a perceived utility u_j. This takes the following additively separable form (see, e.g., Anderson et al., 1992):

$$u_j = v_j + \epsilon_j, \quad j = 1 \ldots J,$$

where v_j is the j-th component of a fixed row vector of systematic utilities $\mathbf{v} = (v_1, \ldots, v_J)$, and ϵ_j is the j-th component of a row vector $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_J)$ of random terms. The systematic utility vector $\mathbf{v}$ is a function of individual's and alternative's attributes. The random vector $\boldsymbol{\epsilon}$ has an absolutely continuous multivariate distribution with finite expectation and full support on $\mathbb{R}^J$.

The individual chooses the alternative that maximizes her perceived utility. Therefore, the probability $P_j(\mathbf{v})$ of choosing alternative j is the probability of the event $\{u_j \geq u_i, \ \forall i \neq j\}$:

$$P_j(\mathbf{v}) = \Pr(u_j \geq u_i, \ \forall i \neq j), \quad j = 1...J, \tag{1}$$

where $\Pr(\cdot)$ is a probability measure.

Different ARUMs are obtained depending on the joint distribution of the random terms. The most frequently used choice models are multinomial logit, nested logit and probit.

If random terms are independently and identically distributed according to a Gumbel cumulative distribution function (c.d.f.) given by (see, e.g., Kotz and Nadarajah, 2000):

$$F_{\epsilon_j}(z) = \exp\left(-e^{-z/\theta}\right), \quad z \in \mathbb{R}, \ j = 1...J,$$

where θ is a positive scale parameter, then we have the multinomial logit model. The probabilities take the following form (see: McFadden, 1974; Train, 2009):

$$P_j(\mathbf{v}) = \frac{e^{v_j/\theta}}{\sum_{k=1}^{J} e^{v_k/\theta}}, \quad j = 1...J.$$

For identification problems, it is usually assumed in estimation that $\theta = 1$.

If the set of alternatives is partitioned into K subsets I_k, $k = 1...K$, and if the alternatives are correlated within subsets and uncorrelated across subsets with multivariate joint c.d.f. of the random terms:

$$F_{\boldsymbol{\epsilon}}(z_1, ..., z_J) = \exp\left[-\sum_{k=1}^{K}\left(\sum_{j \in I_k} e^{-z_j/\theta_k}\right)^{\theta_k}\right], \quad z_1, ..., z_J \in \mathbb{R},$$

where θ_k are parameters satisfying $0 < \theta_k \leq 1$, $k = 1...K$, then we have the two-level nested logit model. The marginals are Gumbel. When $\theta_k = 1$ for all k, the model reduces to multinomial logit.

W.l.o.g., the systematic utility can be decomposed into two parts:

$$v_j = \hat{v}_k + v_{j|k}, \quad j \in I_k, \ k = 1...K,$$

where $\hat{v}_k$ includes only the attributes shared by all the alternatives in the subset I_k, and $v_{j|k}$ includes only the attributes that take different values across alternatives of the subset I_k.

The probability of choosing alternative j belonging to subset I_k takes the form (for derivation see: Ben-Akiva and Lerman, 1985; Train, 2009):

$$P_j(\mathbf{v}) = \frac{e^{\hat{v}_k + \Upsilon_k}}{\sum_{h=1}^{K} e^{\hat{v}_h + \Upsilon_h}} \frac{e^{v_{j|k}/\theta_k}}{\sum_{i \in I_k} e^{v_{i|k}/\theta_k}}, \quad j \in I_k, \ k = 1...K,$$

where Υ_k, related to subset I_k, is the inclusive variable of the lower level choices, given by:

$$\Upsilon_k = \theta_k \ln \sum_{i \in I_k} e^{v_{i|k}/\theta_k}, \quad k = 1...K.$$

If the random vector $\boldsymbol{\epsilon}$ is distributed according to a multivariate normal, then we have the probit model. Probabilities have not not a closed form. They can be obtained by simulation, i.e. drawing from the distribution of the random terms, directly from Eq. (1) (Train, 2009).

2.2. Systematic utility specifications

Let y be individual's income, p_j the price (monetary cost) of alternative j, $\overline{v}_j$ the contribution of other quality attributes of alternative j (e.g. travel time). We consider the following specification:

$$v_j = w_j \left(y - p_j\right) + \overline{v}_j, \quad j = 1...J, \tag{2}$$

where $w_j \left(y - p_j\right)$ is a strictly increasing function of the residual income $y - p_j$.

A theoretical justification of Eq. (2) is provided in Appendix B. We distinguish two cases, one without income effect and one with income effect.

Without income effect, we consider a linear dependence of systematic utilities in residual income $y - p_j$:

$$v_j = \lambda \left(y - p_j\right) + \overline{v}_j, \quad j = 1...J, \tag{3}$$

where $\lambda > 0$ is the constant marginal utility of income. Notice that choice probabilities are independent of income.

With income effect, two specifications are commonly used.
One has alternative-specific constant marginal utility of income:

$$v_j = \lambda_j \left(y - p_j\right) + \overline{v}_j, \quad j = 1...J,$$

with $\lambda_j > 0$, $j = 1...J$.
The other is the translog which is non-linear in income:

$$v_j = \lambda \ln \left(y - p_j\right) + \overline{v}_j, \quad j = 1...J.$$

The translog is appealing because it models a decreasing marginal utility of income, with the value of one dollar decreasing with increasing income.

3. Expectation of compensating variation

3.1. Without income effect

In this section we consider the case of no income effect with systematic utilities of the form in Eq. (3). Consider a change of state: $p_1' \to p_j''$, $\overline{v}_j' \to \overline{v}_j''$, $j = 1...J$. Assume the random terms are unchanged in the two states without and with the change. Then, the random compensating variation cv satisfies:

$$\max_{j=1...J} \left[\lambda \left(y - p_j'\right) + \overline{v}_j' + \epsilon_j\right] = \max_{j=1...J} \left[\lambda \left(y - cv - p_j''\right) + \overline{v}_j'' + \epsilon_j\right].$$

Notice that the compensating variation cv is independent of income. Let the row vector of the simplified systematic utilities be $\hat{\mathbf{v}} = (\hat{v}_1, ..., \hat{v}_J) = (v_1 - \lambda y, ..., v_J - \lambda y)$.

McFadden (1999; p. 258) has proved that the expectation of the random compensating variation is:

$$\mathbb{E}\left[cv\right] = \frac{1}{\lambda}\left\{\mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j'' + \epsilon_j\right)\right] - \mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j' + \epsilon_j\right)\right]\right\}. \tag{4}$$

It is straightforward to extend the proof by McFadden (1999) to the case of changing choice sets. Assume that in the state without the change we have the alternatives $j' = 1, ..., J'$, and in the state with the change the alternatives $j'' = 1, ..., J''$. Then, the expectation of the compensating variation is:

$$\mathbb{E}\left[cv\right] = \frac{1}{\lambda}\left\{\mathbb{E}\left[\max_{j''=1...J''}\left(\hat{v}_{j''}'' + \epsilon_{j''}\right)\right] - \mathbb{E}\left[\max_{j'=1...J'}\left(\hat{v}_{j'}' + \epsilon_{j'}\right)\right]\right\}.$$

Having assumed again an unchanged choice set, we specialize Eq. (4) to multinomial and nested logit. For multinomial logit, the expectation of the maximum utility is (McFadden, 1978):

$$\mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j + \epsilon_j\right)\right] = \ln \sum_{j=1}^{J} e^{\hat{v}_j} + \gamma,$$

where $\gamma \cong 0.57$ is Euler's constant.

Therefore, the expectation of the compensating variation is given by the monetized difference of the logsum:

$$\mathbb{E}\left[cv\right] = \frac{1}{\lambda}\left(\ln \sum_{j=1}^{J} e^{\hat{v}_j''} - \ln \sum_{j=1}^{J} e^{\hat{v}_j'}\right). \tag{5}$$

The estimator of the monetized difference of the logsum in Eq. (5), derived from the maximum likelihood estimators of the coefficients of the systematic utilities, is consistent, asymptotically efficient and asymptotically normal under mild conditions (Delle Site, 2021). Conditions include a strictly positive marginal utility of income λ. The delta method (Mittelhammer, 2013) provides an estimator of the asymptotic variance of the estimator of the monetized difference of the logsum. From this, large sample confidence bounds can be computed.

For the two-level nested logit, the expectation of the maximum utility is (McFadden, 1978):

$$\mathbb{E}\left[\max_{j=1..J}\left(\hat{v}_j+\epsilon_j\right)\right]=\ln\sum_{k=1}^{K}\left(\sum_{j\in I_k}e^{\hat{v}_j/\theta_k}\right)^{\theta_k}+\gamma.$$

Therefore, the expectation of the compensating variation is:

$$\mathbb{E}\left[cv\right]=\frac{1}{\lambda}\left[\ln\sum_{k=1}^{K}\left(\sum_{j\in I_k}e^{\hat{v}''_j/\theta_k}\right)^{\theta_k}-\ln\sum_{k=1}^{K}\left(\sum_{j\in I_k}e^{\hat{v}'_j/\theta_k}\right)^{\theta_k}\right].$$

For any ARUM, the expectation of the compensating variation is also given by the following path-independent line integral (McFadden, 1999):

$$\mathbb{E}\left[cv\right]=\frac{1}{\lambda}\int_{\hat{\mathbf{v}}'}^{\hat{\mathbf{v}}''}\sum_{j=1}^{J}P_j\left(\hat{\mathbf{v}}\right)d\hat{v}_j. \tag{6}$$

where $\hat{\mathbf{v}}'=\left(\hat{v}'_1,...,\hat{v}'_J\right),\hat{\mathbf{v}}''=\left(\hat{v}''_1,...,\hat{v}''_J\right)$.

This integral is important, because it is the starting point to prove the linkage between the expectation of the compensating variation and the rule-of-a-half. Therefore, we provide here a justification of Eq. (6) along the lines of McFadden (1999).

Then, define:

$$\mathit{\Gamma}=\frac{1}{\lambda}\mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j+\epsilon_j\right)\right].$$

Consider the total differential of $\mathit{\Gamma}$ with respect to the systematic utilities:

$$d\mathit{\Gamma}=\frac{1}{\lambda}\sum_{j=1}^{J}\frac{\partial\mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j+\epsilon_j\right)\right]}{\partial\hat{v}_j}d\hat{v}_j. \tag{7}$$

Using a property of ARUMs (the so-called Williams-Daly-Zachary theorem; see McFadden, 1981) we get:

$$\frac{\partial\mathbb{E}\left[\max_{j=1...J}\left(\hat{v}_j+\epsilon_j\right)\right]}{\partial\hat{v}_j}=P_j\left(\hat{\mathbf{v}}\right),\quad j=1...J.$$

By substitution into Eq. (7) we obtain:

$$d\Gamma = \frac{1}{\lambda}\sum_{j=1}^{J} P_j\left(\hat{\mathbf{v}}\right) d\hat{v}_j,$$

which yields, by integration, the line integral of Eq. (6). The integral is path independent because the differential is exact (Galbis and Maestre, 2012).

Figure 1 illustrates the geometry of the line integral in a case of logit with two alternatives, with $\lambda = 1$ and when the path of integration is the line segment ℓ, between $\hat{\mathbf{v}}'$ with coordinates $(-5, -1.5)$ and $\hat{\mathbf{v}}''$ with coordinates $(1, 1.5)$, given by the equation $\hat{v}_2 = 0.5\hat{v} + 1$. It is possible to prove that the line integral is given by the sum of the two shaded areas drawn over the planes $\hat{v}_2 = -1.5$ and $\hat{v}_1 = -5$ (see Delle Site, 2008, for the general case with more than two alternatives).

The areas are given by, respectively, the ordinary integrals $\int_{\hat{v}_1'}^{\hat{v}_1''} \tilde{P}_1(\hat{v}_1)d\hat{v}_1$ and $\int_{\hat{v}_2'}^{\hat{v}_2''} \tilde{P}_2(\hat{v}_2)d\hat{v}_2$, where $\tilde{P}_1(\hat{v}_1)$ (resp. $\tilde{P}_2(\hat{v}_2)$) is the curve $P_1\left(\hat{v}_1, \hat{v}_2\right)$ (resp. $P_2\left(\hat{v}_1, \hat{v}_2\right)$) evaluated over the parameterized line ℓ. The parametric equations of the line ℓ express the systematic utility of one alternative as function of the systematic utility of the other:

$$\hat{v}_k = \hat{v}_k\left(\hat{v}_i\right) = \hat{v}_k' + \left(\hat{v}_i - \hat{v}_i'\right)\frac{\hat{v}_k'' - \hat{v}_k'}{\hat{v}_i'' - \hat{v}_i'}, \quad k = 1, 2,\ i \neq k.$$

As first proved by Jara-Díaz (1990), computation of the line integral in Eq.

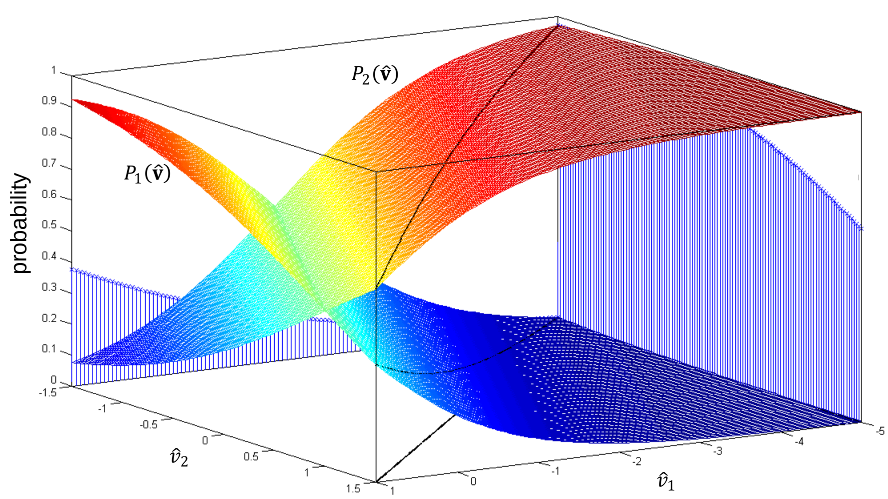

Figure 1: Geometric representation of the line integral of Eq. (6) with two alternatives

(6) over the segment between $\hat{\mathbf{v}}'$ and $\hat{\mathbf{v}}''$ and linearization of demand yields the

following result: the expectation of the compensating variation is approximately given by the following expression, referred to as rule-of-a-half:

$$\mathbb{E}\left[cv\right] \cong \frac{1}{2\lambda}\sum_{j=1}^{J}\left[P_j\left(\hat{\mathbf{v}}'\right) + P_j\left(\hat{\mathbf{v}}''\right)\right]\left(\hat{v}_j'' - \hat{v}_j'\right), \tag{8}$$

where $P_j\left(\hat{\mathbf{v}}'\right)$ and $P_j\left(\hat{\mathbf{v}}''\right)$ are the probabilities in, respectively, the state without and the state with change. Geometrically and with reference to Figure 1, the rule-of-a-half is equivalent to the sum of the two trapezoidal areas obtained from the shaded areas over the planes $\hat{v}_2 = -1.5$ and $\hat{v}_1 = -5$.

The rule-of-a-half lends itself to a nice, but incorrect, interpretation. Rewrite Eq. (8) in the following form:

$$\mathbb{E}\left[cv\right] \cong \frac{1}{\lambda}\sum_{j=1}^{J}P_j\left(\hat{\mathbf{v}}'\right)\left(\hat{v}_j'' - \hat{v}_j'\right) + \frac{1}{2\lambda}\sum_{j=1}^{J}\left[P_j\left(\hat{\mathbf{v}}''\right) - P_j\left(\hat{\mathbf{v}}'\right)\right]\left(\hat{v}_j'' - \hat{v}_j'\right).$$

For each alternative, if $P_j\left(\hat{\mathbf{v}}''\right) > P_j\left(\hat{\mathbf{v}}'\right)$, one would be induced to interpret the first term in the r.h.s. as the welfare change related to the existing demand $P_j\left(\hat{\mathbf{v}}'\right)$, and the second term as the welfare change related to the newly created demand $P_j\left(\hat{\mathbf{v}}''\right) - P_j\left(\hat{\mathbf{v}}'\right)$. However, the interpretation is not correct because it does not consider the probabilistic nature of the model. The rigorous attribution of the welfare change to shifters and non shifters is to be made on the basis of the expectations of the compensating variation conditional on the transitions, as it will be seen in Section 4.

3.2. With income effect

In this section we consider the case of income effect with systematic utilities of the general form in Eq. (2). Consider a change of state: $p_j' \to p_j'', \overline{v}_j' \to \overline{v}_j'', j = 1...J$. Assume the random terms are unchanged in the two states without and with the change. Then, the random compensating variation cv satisfies:

$$\max_{j=1...J}\left[w_j\left(y - p_j'\right) + \overline{v}_j' + \epsilon_j\right] = \max_{j=1...J}\left[w_j\left(y - cv - p_j''\right) + \overline{v}_j'' + \epsilon_j\right]. \tag{9}$$

The expectation of the compensating variation is not available in closed form. It can be computed by simulation (Cherchi et al., 2004; Herriges and Kling, 1999; McFadden, 1999), i.e. by drawing from the distribution of the random terms, or numerically by means of one-dimensional integrals as follows.

One formula for the expectation of the compensating variation is provided by de Palma and Kilani (2011; Theorem 3). They obtain it based on Marshallian, i.e. observed, transition probabilities (see Section 4).

Let $\delta_j = v_j'' - v_j'$, $j = 1...J$, be the changes in the systematic utilities. Assume, without loss of generality, the ordering $\delta_1 \leq ... \leq \delta_J$. We use the notation $x^+ = \max\left(x, 0\right)$. We have the formula:

$$\mathbb{E}\left[cv\right] = \overline{\psi}_J - \sum_{j=1}^{J} \int_{\psi_{jj}}^{\overline{\psi}_J} P_j \left(v_1' + \delta_1^+ \left(c\right), ..., v_J' + \delta_J^+ \left(c\right)\right) dc, \tag{10}$$

where $\overline{\psi}_J$ is given by $\overline{\psi}_J = \max_{j=1,...,J} \psi_{Jj}$; ψ_{Jj} is the compensation of income that satisfies:

$$\delta_j \left(\psi_{Jj}\right) = w_j \left(y - \psi_{Jj} - p_j''\right) + \overline{v}_j'' - w_j \left(y - p_j'\right) - \overline{v}_j' = \left(\delta_j - \delta_J\right)^+;$$

ψ_{jj} is the income compensation for an individual who chooses j without and with the change:

$$w_j \left(y - p_j'\right) + \overline{v}_j' = w_j \left(y - \psi_{jj} - p_j''\right) + \overline{v}_j''; \tag{11}$$

$\delta_j\left(c\right)$ is the difference between utility with change and income compensated by c and utility without the change:

$$\delta_j \left(c\right) = w_j \left(y - c - p_j''\right) + \overline{v}_j'' - w_j \left(y - p_j'\right) - \overline{v}_j'.$$

Notice that ψ_{jj} is unique because of strictly increasing monotonicity of systematic utility in residual income.

Another formula of the expectation of the random compensating variation is by Karlström (2014). Preliminarily, we need to introduce the random expenditure function, because the formula applies the relationship between the expectation of the compensating variation and the expectation of the expenditure function.

We define the random variable maximum utility:

$$\bar{u} \equiv \max_{j=1...J} u_j = \max_{j=1...J} \left(v_j + \epsilon_j\right).$$

The random expenditure function $m\left(\bar{u}'\right)$ necessary to achieve the utility level $\bar{u}'$ of the state without change is defined implicitly by (Karlström and Morey, 2004; Dagsvik and Karlström, 2005):

$$\bar{u}' = \max_{j=1...J} \left[w_j \left(m\left(\bar{u}'\right) - p_j''\right) + \overline{v}_j'' + \epsilon_j\right]. \tag{12}$$

Consider the expenditure $m_{\to j}\left(\bar{u}'\right)$, conditional on the choice of alternative j in the state with change, necessary to achieve the utility level $\bar{u}'$ of the state without change. The conditional expenditure $m_{\to j}\left(\bar{u}'\right)$ satisfies:

$$\bar{u}' = w_j \left(m_{\to j} \left(\bar{u}' \right) - p''_j \right) + \overline{v}''_j + \epsilon_j, \quad j = 1...J. \tag{13}$$

Due to the increasing monotonicity of systematic utilities in income, having taken into account Eq. (12) and Eq. (13), we have:

$$m \left(\bar{u}' \right) = \min_{j=1...J} \left[m_{\to j} \left(\bar{u}' \right) \right]. \tag{14}$$

The compensating variation $cv_{\to j}$ conditional on the choice of alternative j in the state with change satisfies:

$$\bar{u}' = w_j \left(y - cv_{\to j} - p''_j \right) + \overline{v}''_j + \epsilon_j, \quad j = 1...J. \tag{15}$$

Due to the increasing monotonicity of the systematic utilities in income, having taken into account Eq. (9) and Eq. (15), we have:

$$cv = \max_{j=1...J} cv_{\to j}. \tag{16}$$

By definition of conditional expenditure and conditional compensating variation, we have:

$$m_{\to j} \left(\bar{u}' \right) = y - cv_{\to j}, \quad j = 1...J.$$

Therefore, by Eq. (14) and Eq. (16) we get:

$$m \left(\bar{u}' \right) \; = \; \min_{j=1...J} m_{\to j} \left(\bar{u}' \right) \; = \; \min_{j=1...J} \left(y - cv_{\to j} \right) \; = \; y - \max_{j=1...J} cv_{\to j} \; = \; y - cv.$$

By taking expectations we get:

$$\mathbb{E} \left[m \left(\bar{u}' \right) \right] = y - \mathbb{E} \left[cv \right]. \tag{17}$$

Karlström (2014) obtains the expression of $\mathbb{E} \left[m \left(\bar{u}' \right) \right]$, from which, based on Eq. (17), he gets:

$$\mathbb{E} \left[cv \right] = y - \sum_{j=1}^{J} \int_0^{\mu_{jj}} P_j \left(g_1 \left(m \right), ..., g_J \left(m \right) \right) dm, \tag{18}$$

where μ_{jj} is the expenditure needed with the change to restore the without change level of utility if alternative j is chosen without and with the change:

$$w_j\left(y-p_j'\right)+\overline{v}_j'=w_j\left(\mu_{jj}-p_j''\right)+\overline{v}_j'',\quad j=1...J,$$

and

$$g_j\left(m\right)=\max\left[w_j\left(y-p_j'\right)+\overline{v}_j',w_j\left(m-p_j''\right)+\overline{v}_j''\right],\quad j=1...J.$$

Notice that μ_{jj} is unique because of strictly increasing monotonicity of systematic utility in residual income.

Both formulas of Eq. (10) by de Palma and Kilani (2011) and Eq. (18) by Karlström (2014) generally require numerical integration, because the indefinite integrals in the formulas cannot be expressed in closed form. The integrand functions are not in a closed form, because they require the computation of a maximum function. Libraries implemented in main software environments allow today easy computation of these integrals.

3.3. Changing the random terms

In the two previous sections, the assumption of unchanged random terms is made. Here, we consider the changes in the random terms: $\epsilon_j' \to \epsilon_j''$, $j=1,...,J$. Delle Site and Salucci (2013, Proposition 4) prove that, without income effect, the expectation of the compensating variation is:

$$\mathbb{E}\left[cv\right]=\frac{1}{\lambda}\left\{\mathbb{E}\left[\max_{j=1...J}\left(v_j''+\epsilon_j''\right)\right]-\mathbb{E}\left[\max_{j=,...J}\left(v_j'+\epsilon_j'\right)\right]\right\},$$

which depends only on the marginal distributions of the random terms without change and with change. This theoretical finding is not trivial because ϵ_j' is correlated with ϵ_j''. The property was first illustrated by numerical experiments by Zhao et al. (2012).

With income effect, the expectation of the compensating variation can be computed by simulation (see examples in Delle Site and Salucci, 2013).

4. Transition choice probabilities and conditional expectations of compensating variation

4.1. Transition choice probabilities

In this section, we provide expressions for the transition probabilities, i.e. probabilities of choosing alternative i without the change and alternative j with the change. Transition probabilities are of interest because they provide, for a population of homogenous individuals (in terms of systematic utilities), the estimation of the shares of those who stay on the same alternative (non shifters) and of those who change alternative (shifters). In the case without income effect, income is irrelevant to transition probability computation.

Consider a change of state: $\mathbf{v}' = (v'_1, ..., v'_J) \to \mathbf{v}'' = (v''_1, ..., v''_J)$. Random terms are assumed unchanged. Again, let $\delta_j = v''_j - v'_j$, $j = 1...J$, and assume the ordering $\delta_1 \leq ... \leq \delta_J$. de Palma and Kilani (2011, Theorem 1) prove that the probability $P_{i\to j}(\mathbf{v}', \mathbf{v}'')$ of the transition from alternative i to j is:

$$P_{i\to j}(\mathbf{v}', \mathbf{v}'') = \begin{cases} P_i\left(v'_1 + (\delta_1 - \delta_i)^+, ..., v'_J + (\delta_J - \delta_i)^+\right), & j = i; \\ \int_{\delta_i}^{\delta_j} \Pi_i^j\left(v'_1 + (\delta_1 - z)^+, ..., v'_J + (\delta_J - z)^+\right) dz, & j > i; \\ 0, & j < i, \end{cases}$$

where $x^+ = \max(x, 0)$ and $\Pi_i^j(\zeta_1, ..., \zeta_J) = -\partial P_i(\zeta_1, ..., \zeta_J)/\partial \zeta_j$.

For multinomial logit, transition probabilities are available in closed form (de Palma and Kilani, 2011, Proposition 1):

$$P_{i\to j}(\mathbf{v}', \mathbf{v}'') = \begin{cases} \frac{e^{v'_i}}{\Omega_i}, & j = i; \\ \sum_{r=i}^{j-1}\left(\frac{e^{v'_i}}{\Omega_{r+1}} - \frac{e^{v'_i}}{\Omega_r}\right)\frac{e^{v''_j}}{\sigma_r}, & j > i; \\ 0, & j < i, \end{cases} \tag{19}$$

where:

$$\begin{aligned} \sigma_r &= \sigma_0 - \sum_{k\leq r} e^{v''_k}, \quad r = 1, ..., J; \\ \sigma_0 &= \sum_k e^{v''_k}; \\ \Omega_r &= s_r + \sigma_r \cdot e^{-\delta_r}, \quad r = 1, ..., J; \\ s_r &= \sum_{k\leq r} e^{v'_k}, \quad r = 1, ..., J. \end{aligned}$$

4.2. Conditional expectations of compensating variation

In this section, we provide expressions for the expectations of the compensating variation conditional on the transitions. Conditional expectations of the compensating variation are of interest because they provide, for a population of homogenous individuals (in terms of systematic utilities), the estimation of the average welfare change for the sub-populations of shifters and non shifters.

Random terms are assumed unchanged. Hereafter, only transitions $i \to j$ for which we have a strictly positive probability $P_{i\to j} > 0$ are considered. De Palma and Kilani (2011, Theorem 3) prove that the expectation of the random compensating variation conditional on the transition $i \to j$ is:

$$\mathbb{E}[cv_{i\to j}] =$$
$$\overline{\psi}_j - \frac{1}{P_{i\to j}(\mathbf{v}', \mathbf{v}'')}\int_{\underline{\psi}_{ij}}^{\overline{\psi}_j} P_{i\to j}\left(v'_1 + \delta_1^+(c), ..., v'_J + \delta_J^+(c); \mathbf{v}''\right) dc, \quad j \geq i,$$

where $\overline{\psi}_j = \max_{i=1...J} \psi_{ji}$, $\overline{\psi}_{ij} = \max(\psi_{ii}, \psi_{ij})$, $\delta_j(c) = w_j\left(y - c - p''_j\right) + \overline{v}''_j - v'_j$, ψ_{ij} satisfies $\delta_j(\psi_{ij}) = (\delta_j - \delta_i)^+$.

For multinomial logit, in the case without income effect of specification (3), the expectation of the compensating variation conditional on the transition $i \to j$ is available in closed form (de Palma and Kilani, 2011, proposition 4):

$$\mathbb{E}\left[cv_{i\to j}\right] = \begin{cases} \psi_{ii}, & j = i; \\ \frac{1}{\Xi_{ij}} \sum_{r=i}^{j-1} \frac{1}{\sigma_r} \left(\frac{\psi_{(r+1)(r+1)}}{\Omega_{r+1}} - \frac{\psi_{rr}}{\Omega_r} - \tau_r \right), & j > i, \end{cases} \tag{20}$$

where

$$\Xi_{ij} = \sum_{r=i}^{j-1} \frac{1}{\sigma_r} \left(\frac{1}{\Omega_{r+1}} - \frac{1}{\Omega_r} \right), \quad j > i,$$

and

$$\tau_r = \frac{1}{\lambda} \left(\frac{\delta_{r+1} - \delta_r + \ln \Omega_{r+1} - \ln \Omega_r}{s_r} \right), \quad r = 1...J - 1.$$

Notice that, since there is no income effect, σ_0 and σ_r, s_r, $r = 1, ..., J$, are obtained using the vectors $\hat{\mathbf{v}}'$ and $\hat{\mathbf{v}}''$ of the simplified systematic utilities.

By the law of total expectation, the expectation of the compensating variation conditional on the choice of alternative i without the change is:

$$\mathbb{E}\left[cv_{i\to}\right] = \frac{1}{P_i(\mathbf{v}')} \sum_{j\geq i} P_{i\to j}(\mathbf{v}', \mathbf{v}'') \, \mathbb{E}\left[cv_{i\to j}\right], \quad i = 1...J, \tag{21}$$

while the expectation of the compensating variation conditional on the choice of alternative j with the change is:

$$\mathbb{E}\left[cv_{\to j}\right] = \frac{1}{P_j(\mathbf{v}'')} \sum_{i\leq j} P_{i\to j}(\mathbf{v}', \mathbf{v}'') \, \mathbb{E}\left[cv_{i\to j}\right], \quad j = 1...J. \tag{22}$$

An additional application of the law of total expectation provides the unconditional expectation of the compensating variation:

$$\mathbb{E}[cv] = \sum_{i=1}^{J} P_i(\mathbf{v}') \, \mathbb{E}\left[cv_{i\to}\right] = \sum_{j=1}^{J} P_j(\mathbf{v}'') \, \mathbb{E}\left[cv_{\to j}\right].$$

5. Distribution of compensating variation and inequality analysis

5.1. Distribution of compensating variation

The c.d.f. $\Phi_{j\to}$ of the compensating variation $cv_{j\to}$ conditional on the choice of alternative j without change is given by the following conditional probability:

$$\Phi_{j\to}(c) = \Pr\left(cv_{j\to} \leq c \mid v'_j + \epsilon_j = \max_{k=1...J} (v'_k + \epsilon_k)\right), \quad j = 1...J.$$

The support and the expression of the conditional compensating variation $cv_{j\rightarrow}$, and the support and the expression of the unconditional compensating variation cv are available from de Palma and Kilani (2011). Random terms are assumed unchanged.

The conditional compensating variation $cv_{j\rightarrow}$ has support $\left[\psi_{jj},\hat{\psi}\right]$, where ψ_{jj} is the income compensation that satisfies Eq. (11), and $\hat{\psi}$ is given by $\hat{\psi}=\max_{k=1...J}\left(\psi_{kk}\right)$. The unconditional compensating variation cv has support given by $\left[\min_{k=1...J}\left(\psi_{kk}\right),\hat{\psi}\right]$.

Define the vector:

$$\mathbf{v}\left(c\right)=\left(\max\left[v_1',w_1\left(y-c-p_1''\right)+\overline{v}_1''\right],...,\max\left[v_J',w_J\left(y-c-p_J''\right)+\overline{v}_J''\right]\right).$$

The conditional compensating variation $cv_{j\rightarrow}$ has, for any c of its support, c.d.f.:

$$\Phi_{j\rightarrow}\left(c\right)=\frac{P_j\left(\mathbf{v}\left(c\right)\right)}{P_j\left(\mathbf{v}'\right)},\quad j=1...J. \tag{23}$$

The unconditional compensating variation cv has, for any c of its support, c.d.f.:

$$\Phi\left(c\right)=\sum_{j=1}^{J}H_j\left(c\right)P_j\left(\mathbf{v}\left(c\right)\right), \tag{24}$$

where $H_j\left(c\right)=1$ if $c\geq\psi_{jj}$, $H_j\left(c\right)=0$ otherwise (Heaviside function at ψ_{jj}). This distribution is discrete-continuous.

The shares of losers and of winners are provided by, respectively, $\Phi\left(c=0\right)$ and $1-\Phi\left(c=0\right)$.

5.2. Inequality analysis

The analysis of inequality in the distribution of the compensating variation can be carried out based on the Lorenz curve and the Gini coefficient. The Lorenz curve can be used to relate the cumulative proportion of users to the cumulative proportion of welfare change received (random compensating variation) when users are arranged in ascending order of their welfare change. The associated Gini coefficient is a synthetic index of inequality (see, e.g., Gastwirth, 1972). A Gini coefficient of value zero expresses perfect equality, of value one maximal inequality.

The common use of the Lorenz curve and Gini coefficient is for representation of situations with positive wealth only. In this occurrence, i.e., only gains in terms of random compensating variation, as illustrated in Figure 2, the Lorenz curve is included in the interval $[0,1]$. The Gini coefficient is defined as the ratio of the area between the Lorenz curve and the $+45°$ line (area A) to the area delimited by the the $+45°$ line and the axes (area A+B). The area A is called the area of concentration. The $+45°$ line represents complete equality in gains: everyone has the same gain.

Extension to the case of losses only is straightforward: as illustrated in Figure 3, the Lorenz curve is included in the interval $[-1, 0]$, and the Gini coefficient is defined as the ratio of area A to area A+B. The $-45°$ line represents complete equality in losses: everyone has the same loss. The Gini coefficient is, in both cases, included in the interval between 0 (complete equality) and 1 (complete inequality, i.e. the condition where the marginal user takes the full gain, or loss, and the others take none).

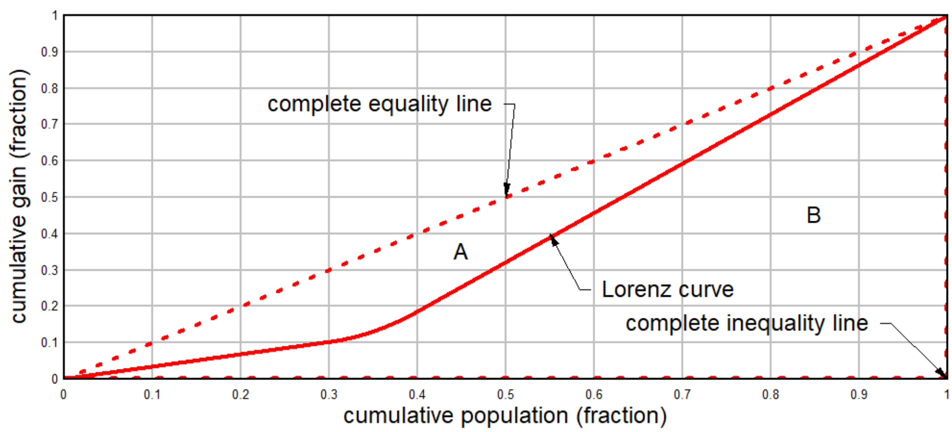

Figure 2: A Lorenz curve in the case of gains only

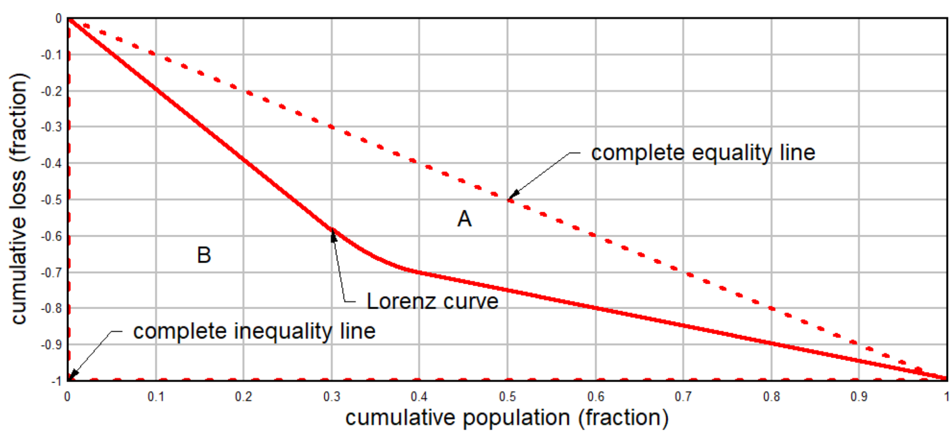

Figure 3: A Lorenz curve in the case of losses only

Extension to mixed cases with both positive and negative wealth has been dealt with by a few authors (see, e.g., Schutz, 1951, and Chen et al., 1982; De Battisti et al., 2019, provide a state of the art on the topic). In the literature, only the case with positive expectation of wealth is considered. Extended

Lorenz curves and adjusted Gini coefficients have been proposed. However, intepretation in terms of inequality is controversial. For this reason, in the case where the compensating variation takes both negative and positive values, we propose a distinct inequality analysis for losses and for gains. To this aim, we define the c.d.f. $\Phi^{-}(c)$ of the non-gain values of the compensating variation cv:

$$\Phi^{-}(c) = \frac{\Phi(c)}{\Phi(0)}, \quad \Phi(0) \neq 0, c \leq 0,$$

and the c.d.f. $\Phi^{+}(c)$ of the non-loss values of the compensating variation cv:

$$\Phi^{+}(c) = \frac{\Phi(c) - \Phi(0)}{1 - \Phi(0)}, \quad \Phi(0) \neq 1, c \geq 0.$$

Evidently, in the case of non-gains only, we get $\Phi^{-}(c) = \Phi(c)$ since $\Phi(0) = 1$. In the case of gains only, we get $\Phi^{+}(c) = \Phi(c)$ since $\Phi(0) = 0$.

Assume that the population of users is in ascending order of relative value of compensating variation.

Proposition 1. *The Lorenz curve $\mathcal{L}$, respectively for non-gains $cv \leq 0$ and for non-losses $cv \geq 0$, is:*

$$\mathcal{L}(\pi) = \begin{cases} \frac{1}{|\mathbb{E}^{-}[cv]|} \int_0^{\pi} \Phi^{-^{-1}}(t)\, dt, \ cv \leq 0, \ 0 \leq \pi \leq 1; \\ \frac{1}{\mathbb{E}^{+}[cv]} \int_0^{\pi} \Phi^{+^{-1}}(t)\, dt, \ cv \geq 0, \ 0 \leq \pi \leq 1, \end{cases}$$

where π is the cumulative proportion of the population of users, $\Phi^{-^{-1}}(\cdot)$ and $\Phi^{+^{-1}}(\cdot)$ are, respectively, the inverse of the c.d.f. $\Phi^{-}(c)$ of the non-gain values of cv and the inverse of the c.d.f. $\Phi^{+}(c)$ of the non-loss values of cv, $\mathbb{E}^{-}[cv]$ and $\mathbb{E}^{+}[cv]$ are, respectively, the expectation of the non-positive values of cv and the expectation of the non-negative values of cv, $|\cdot|$ denotes the absolute value.

Proof. From Lorenz curve definition (Gastwirth, 1972). □

The inversion of $\Phi^{-}(c)$ and $\Phi^{+}(c)$, and subsequent integration required to draw the Lorenz curve can be carried out numerically.

Proposition 2. *The Gini coefficient G can be obtained, respectively for non-gains $cv \leq 0$ and for non-losses $cv \geq 0$, by simulation using the formulas :*

$$G = \begin{cases} \frac{1}{2|\mathbb{E}_{-}[cv]|} \mathbb{E}\left[|cv_1 - cv_2|\right], \ cv_1, cv_2 \leq 0; \\ \frac{1}{2\mathbb{E}_{+}[cv]} \mathbb{E}\left[|cv_1 - cv_2|\right], \ cv_1, cv_2 \geq 0, \end{cases}$$

where cv_1 and cv_2 are independent copies of cv.

Proof. Based on Kendall and Stuart (1958). □

6. Observed and unobserved heterogeneity

Assume that the individuals are heterogeneous with respect to attributes of the systematic utilities (observed heterogeneity), or with respect to estimation coefficients of the systematic utilities (unobserved heterogeneity). One may wish to carry out welfare analysis for the aggregate of the individuals.

In transportation applications, the case of observed heterogeneity is the one where we have heterogeneous individual attributes, such as income, or heterogeneous alternative attributes, the latter situation typically arising when individuals of different origin-destination pairs are considered. The case of unobserved heterogeneity is, as an example, the one where a distribution of the time and cost coefficients is accounted for. The latter gives rise to the popular mixed, or random coefficients, logit (see, among the others, Train, 2009).

6.1. Distribution of compensating variation

In both the observed and the unobserved case, it is possible to derive an aggregate c.d.f. of the compensating variation on the basis of the law of total probability.

Proposition 3. *In the observed heterogeneity case, if there are N sample individuals, each with a weight ω_n in the population and a c.d.f. of the compensating variation $\Phi_n(c)$, the aggregate c.d.f. is:*

$$\Phi(c) = \frac{1}{\sum_{n=1}^{N} \omega_n} \sum_{n=1}^{N} \omega_n \Phi_n(c), \quad c \in \mathbb{R}.$$

Proposition 4. *In the unobserved heterogeneity case, consider mixed logit. If $\boldsymbol{\beta}$ is the vector of estimation coefficients of the systematic utilities and $\Phi(\boldsymbol{\beta}; c)$ the c.d.f. of the compensating variation conditional on the values of the coefficients $\boldsymbol{\beta}$, then the unconditional c.d.f. is:*

$$\Phi(c) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \Phi(\boldsymbol{\beta}; c) f(\boldsymbol{\beta}) \, d\boldsymbol{\beta}, \quad c \in \mathbb{R},$$

where $f(\boldsymbol{\beta})$ is the probability density function (p.d.f.) of the coefficients $\boldsymbol{\beta}$.

6.2. Expectation of compensating variation

Assume then that one wishes to consider the expectation of the compensating variation over the aggregate of the individuals.

Proposition 5. *In the observed heterogeneity case, given the expectation $\mathbb{E}_n[cv]$ for sample individual n, the expectation over the aggregate is:*

$$\mathbb{E}[cv] = \frac{1}{\sum_{n=1}^{N} \omega_n} \sum_{n=1}^{N} \omega_n \mathbb{E}_n [cv] .$$

Proof. By the law of total expectation. □

Proposition 6. *In the unobserved heterogeneity case, consider mixed logit. Let* $\mathbb{E}[\boldsymbol{\beta}; c] = \int_{-\infty}^{\infty} c\varphi(\boldsymbol{\beta}; c)\, dc$ *be the expectation of the compensating variation conditional on the values of the coefficients* $\boldsymbol{\beta}$*, where* $\varphi(\boldsymbol{\beta}; c)$ *is the p.d.f. of the compensating variation conditional on* $\boldsymbol{\beta}$*. The unconditional expectation, i.e. the expectation over the aggregate, is:.*

$$E[cv] = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \mathbb{E}[\boldsymbol{\beta}; c]\, f(\boldsymbol{\beta})\, d\boldsymbol{\beta}.$$

Proof. By the law of total probability, the unconditional p.d.f. of the compensating variation is:

$$\varphi(c) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \varphi(\boldsymbol{\beta}; c)\, f(\boldsymbol{\beta})\, d\boldsymbol{\beta}.$$

We have:

$$E[cv] = \int_{-\infty}^{\infty} c\varphi(c)\, dc.$$

The proposition follows from application of the Fubini theorem, by which change of the order of integration is allowed (Billingsley, 1995). □

7. Numerical illustration

7.1. Data

We consider five transportation modes: car (alternative 1), cycling (alternative 2), motorcycle (alternative 3), public transportation (alternative 4), and walking (alternative 5). Systematic utilities are as follows (no income effect):

$$\hat{v}_j = \beta t_j + ASC_j, \quad j = 1, ..., 5,$$

where t_j is travel time on alternative j, β and ASC_j, $j = 1, ..., 5$ are estimation coefficients (ASC stands for alternative specific constant).

A multinomial logit is estimated on the basis of big data from the Rhône Department in East-Central France, having Lyon as capital. Census data related to more than two hundred thousand households surveyed between 2015 and 2019 are used. Surveys have been conducted by INSEE (Institut National de la Statistique et des Études Économiques). The road network for cars, bicycles, motorcycles and pedestrians is retrieved from OpenStreetMap. Travel times are based on the fastest route computed from the open-source routing engine GraphHopper. The fastest route on the public transportation network is computed using timetables retrieved from GTFS (General Transit Feed Specification) files. Additional details and data statistics are found in Javaudin et al. (2021). The estimates of the coefficients of the logit model with the associated t-stat are in Table 1. We have set $ASC_1 = 0$. The econometric software NLogit has been used.

Table 1: Coefficient estimates and t-stat

coeff.	attribute	estimate (t-stat)
β	travel time (hours)	-1.73093 (-163.1)
ASC_2	cycling	-2.11080 (-194.1)
ASC_3	motorcycle	-4.16263 (-230.9)
ASC_4	public transportation	.39506 (51.3)
ASC_5	walking	.13518 (12.3)
221,571	observations	

7.2. Policy

Assume an ideal corridor with commuters travelling towards a central business district. Assume that in the state without policy we have the following travel times: $t_1 = 0.48$ hours on car, $t_2 = 0.58$ hours cycling, $t_3 = 0.18$ hours on motorcycle, $t_4 = 1.18$ hours on public transportation, and $t_5 = 3.3$ hours walking. The associated modal shares are $P_1' = 63.3\%$ for cars, $P_2' = 6.5\%$ for cycling, $P_3' = 1.7\%$ for motorcycle, $P_4' = 28.0\%$ for public transportation, and $P_5' = 0.5\%$ for walking.

A congestion charging policy is considered. The policy changes modal demand. Travel time on car is considered endogenous because it changes with demand in the light of congestion effects, while travel times on the other modes are assumed constant. For illustration aims, the following volume-delay function (hours) is assumed: $t_1 = 0.1372 + 0.686\,(P_1)^{1.5}$.

According to the classical approach, the optimal congestion charge results from the solution of the system optimum problem (Yang and Huang, 2005). In our multi-modal setting, this is the problem where we search for the modal shares that maximize the direct utility of the representative consumer. The direct utility is provided in Anderson et al. (1988). Direct utility is needed because the demand shares are the decision variables of the optimization problem. The problem is subject to the demand conservation and non-negativity constraints. Therefore:

$$\max_{P_i, i=1,...,5} \hat{v}_1 \left[t_1 \left(P_1 \right) \right] P_1 + \sum_{i=2}^{5} \hat{v}_i P_i - \sum_{i=1}^{5} P_i \ln P_i,$$

$$\text{s.t.} \quad \sum_{i=1}^{5} P_i = 1, \quad P_i \geq 0, i = 1, ..., 5.$$

The system optimum modal shares, which are assumed as shares in the state with policy, are $P_1'' = 51.3\%$ for cars, $P_2'' = 8.6\%$ for cycling, $P_3'' = 2.2\%$ for motorcycle, $P_4'' = 37.2\%$ for public transportation, and $P_5'' = 0.7\%$ for walking. For computation, the solvers of the Python Scipy.Optimize library have been used.

The system optimum shares can be obtained in a decentralized way with a charge. From the first-order conditions of the system optimum problem, we obtain that the optimal charge is equal to the congestion externality. The charge η in utility units is:

$$\eta = P_1'' \frac{\partial \hat{v}_1 \left[t_1 \left(P_1'' \right) \right]}{\partial P_1} = -0.652.$$

Having considered an average value of time equal to 9.17 EUR/h (French value from the meta-analysis in Wardman et al., 2012), we obtain a marginal utility of income $\lambda = 0.18876$, and, consequently, a value of the optimal charge of 3.456 EUR/h. The congestion charging policy results in a significant reduction of the car share from 63.3% to 51.3%, with an associated decrease of travel time on this mode from 0.48 to 0.39 hours (19% reduction). Demand on all other modes increase, particularly on public transportation.

7.3. Users' benefits

Table 2 shows the values of transition probabilities (from the alternative on the row to the alternative on the column; based on Eq. 19), and of probabilities in the state without policy (first column on the right) and in the state with policy (bottom row). The values on the main diagonal represents the shares who stay on the same alternative. The only transitions are from car to other modes. This is because car is the only alternative that undergoes a change in systematic utility, for the combined effect of the congestion charge and the changed travel time.

Table 3 shows the values of the average benefit for the full population, for the sub-populations of shifters and non-shifters (based on Eq. 20), for the sub-populations of those who choose a given alternative in the state without policy (based on Eq. 21) and in the state with policy (based on Eq. 22). Since we have no income effect, the average benefit for the full population is given by the monetized logsum of Eq. (5) and equals -1.500 EUR/trip. The rule-of-a-half of Eq. (8) yields -1.498 EUR/trip.

Table 2: Transition probabilities (%)

	1	2	3	4	5	without
1	51.3	2.1	0.5	9.2	0.2	63.3
2	0.0	6.5	0.0	0.0	0.0	6.5
3	0.0	0.0	1.7	0.0	0.0	1.7
4	0.0	0.0	0.0	28.0	0.0	28.0
5	0.0	0.0	0.0	0.0	0.5	0.5
with	51.3	8.6	2.2	37.2	0.7	100.0

Table 3: Average benefit of the charging policy (EUR/trip)

	1	2	3	4	5	without
1	-2.615	-1.323	-1.323	-1.323	-1.323	-2.370
2	-	0.000	-	-	-	0.000
3	-	-	0.000	-	-	0.000
4	-	-	-	0.000	-	0.000
5	-	-	-	-	0.000	0.000
with	-2.615	-0.326	-0.326	-0.326	-0.326	-1.500

Using the results in Section 5, we constructed the north-west and north-east charts in Figure 4. The north-west chart shows the c.d.f. of $cv_{1\rightarrow}$, the compensating variation of those who choose alternative 1 in the state without. The north-east chart shows the c.d.f. of $cv_{2\rightarrow}$, $cv_{3\rightarrow}$, $cv_{4\rightarrow}$ and $cv_{5\rightarrow}$. In all cases we have a step function. For $cv_{1\rightarrow}$, the c.d.f. is 0 for values $< \psi_{11} = -2.615$. At the abscissa -2.615 there is a step. The probability that the compensating variation $cv_{1\rightarrow}$ takes exactly the value $\psi_{11} = -2.615$ is $\Phi_{1\rightarrow}(\psi_{11}) = 81.1\%$. Indeed, the c.d.f. is continuous from the right. As to $cv_{2\rightarrow}$, $cv_{3\rightarrow}$, $cv_{4\rightarrow}$ and $cv_{5\rightarrow}$, since their support is the point $\psi_{ii} = 0$, $i = 2, ..., 5$, their c.d.f. reduces to a point of abscissa equal to 0 and ordinate equal to 1. The charts are based on Eq. 23. The c.d.f. of the unconditional compensating variation cv is shown in the south-west chart of Figure 4. Since the support of the c.d.f. of cv is $-2.615 \leq cv \leq 0$, there are only losers from the policy. The chart is based on Eq. 24.

The Lorenz curve for the distribution of the unconditional compensating variation is shown in the south-east chart of Figure 4. The chart shows an estimate of the Lorenz curve, obtained analytically by linearizing the curved tract of the distribution of the compensating variation. This is based on Proposition 1. The Gini coefficient associated with the random compensating variation, computed by simulation with one million draws, is 0.424. This is based on Proposition 2.

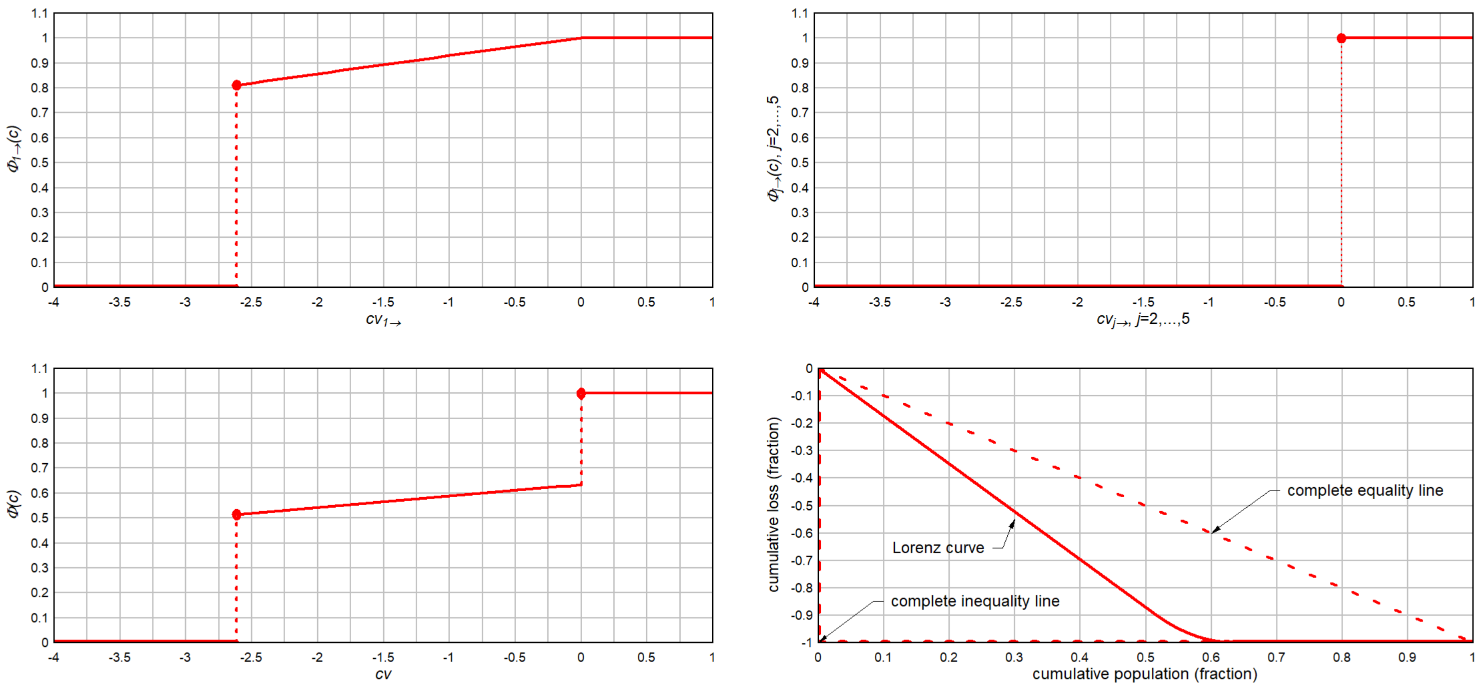

Figure 4: c.d.f. of $cv_{1\rightarrow}$ (north-west), $cv_{2\rightarrow}, cv_{3\rightarrow}, cv_{4\rightarrow}$, and $cv_{5\rightarrow}$(north-east), cv (south-west), and Lorenz curve of cv (south-east) for the illustrative example

8. Conclusion

The apparatus developed in microeconomics for the measurement of welfare in the divisible goods case has been adapted to discrete choice models. The approach to welfare change measurement that is considered mainstream, because of the consistency with the theory of random utility, is based on the expectation of the compensating variation. For multinomial logit without income effect, the expectation of the compensating variation reduces to the popular monetized logsum, which proves to be a robust formula because it holds even for changing random terms. The paper has shown, with an illustrative application, that the recent theoretical developments in welfare measurement for ARUMs can improve the practice of assessment of transportation users' benefits.

The computation of the expectation of the compensating variation by numerical integration is possible today also with income effect, without the need for computationally demanding simulations. The problem of attributing benefits to shifters and non shifters is used to be treated on the basis of the rule-of-a-half. This considers a conventional attribution of benefits to distinct alternatives as well as to non shifters and to new and lost demand. Today, a rigorous solution to the benefit attribution problem, based on random utility theory and using numerical integration without simulation, is available. The approach is able to compute exactly the shifts in demand among pairs of alternatives and to attribute the respective welfare change. It is applicable to both cases without and with income effect, while the rule-of-a-half applies only without income effect.

Knowledge of the distribution of the compensating variation allows the estimation of the shares of winners and losers, and makes inequality analysis possible using numerical methods to compute the Lorenz curve and simulation to compute the Gini coefficient. The application of inequality analysis in the mixed case of both negative and positive welfare change has been presented in the paper. The approach is based on distinct analyses for losses and for gains. A different approach to inequality analysis considers inequality with respect to the sum of income and the equivalent variation, rather than to the compensating variation only. If the welfare change is small enough then the sum of income and the equivalent variation is positive and the conventional definitions, with positive wealth, of Lorenz curve and Gini coefficient apply. Research that adopts this approach and provides the Lorenz curve and Gini coefficient analytically for general ARUMs is ongoing (de Palma et al., 2021).

The welfare inequality analysis has been a first new theoretical contribution of the paper. The second one has been related to the aggregate welfare analysis in the presence of observed and unobserved heterogeneity. These and the other theoretical findings from literature can be implemented in a relatively easy way in software packages that may assist the researcher and the practitioner in project and policy assessment tasks. As seen, the random compensating variation is given prominence in the literature. Without income effect, the expectation of the compensating variation equals the expectation of the equivalent variation (Delle Site, 2021). Further findings related to the equivalent variation, similar to those existent for the compensating variation, are an objective for future research.

Appendix A. Classical microeconomics

This Appendix deals with the microeconomics of divisible goods and is based on Mas-Colell et al. (1995) and Takayama (1994). The Appendix reports on main results. Proofs are found in the two references. Additional references are added where relevant.

Consider the behaviour of a consumer who is assumed to maximize a function of the quantities of M goods subject to a budget constraint:

$$\max_{\mathbf{x}} U(\mathbf{x}),$$
$$\text{s.t.} \sum_{k=0}^{M} r_k x_k \leq y,$$

where $U(\cdot)$ is the direct utility, $\mathbf{x} = (x_0, x_1, ..., x_M)$ is the row vector of quantities of goods, $\mathbf{r} = (r_0, r_1, ..., r_M)$ is the row vector of the prices of goods and y is income. Prices and income are deflated by the price r_0 of the numéraire. If we assume $r_0 = 1$, then x_0 can be interpreted as expenditure on all other goods in the economy. If preferences are locally non satiated (for any bundle of goods there is always another bundle of goods arbitrarily close that is preferred to it, a condition implied by monotonicity of preferences), the budget constraint is satisfied as equality.

We define indirect utility the function of prices and income:

$$V(\mathbf{r}, y) = \max\left\{U(\mathbf{x}) : \sum_{k=0}^{M} r_k x_k = y\right\},$$

and (Marshallian) demands the functions of prices and income:

$$\mathbf{x}(\mathbf{r}, y) = \arg_{\mathbf{x}} \max\left\{U(\mathbf{x}) : \sum_{k=0}^{M} r_k x_k = y\right\}.$$

The indirect utility is continuous in prices and income, non increasing in prices, non decreasing in income, homogenous of degree zero in prices and income (no money illusion), quasi-convex in prices and income. Roy's identity establishes the relationship between indirect utility and demand functions:

$$x_k(\mathbf{r}, y) = -\frac{\partial V(\mathbf{r}, y)/\partial r_k}{\partial V(\mathbf{r}, y)/\partial y}, \quad k = 0, 1...M.$$

Consider now a change of price of one good only: $r'_k \to r''_k$, $k \neq 0$. We define variation of consumer's surplus:

$$\triangle S\left(r'_k, r''_k\right) = -\int_{r'_k}^{r''_k} x_k\left(\mathbf{r}, y\right) dr_k. \tag{25}$$

In this case, consumer's surplus has a nice interpretation which is based on Figure 5. Consider the inverse of the demand function, i.e. the demand price, or, in other words, the maximum price the consumer is willing to pay for a given quantity. Surplus is the difference between the total willingness-to-pay and what is actually paid. The integral in Eq. (25) represents the variation in this difference when price changes.

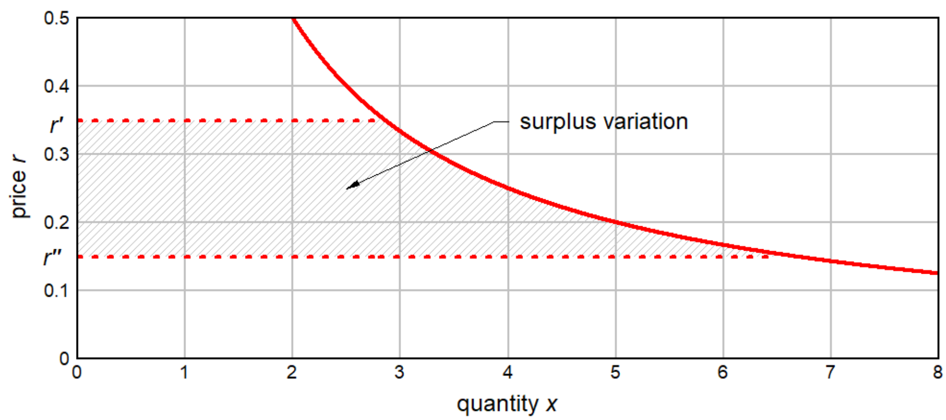

Figure 5: Surplus variation: change in price of one good only

Consider that a change occurs in the price of all goods except the numéraire: $\overline{\mathbf{r}}' = (r'_1, ..., r'_M) \to \overline{\mathbf{r}}'' = (r''_1, ..., r''_M)$, while income is unchanged. The variation of consumer's surplus is defined by the so-called Hotelling integral (Hotelling, 1938):

$$\triangle S\left(\overline{\mathbf{r}}', \overline{\mathbf{r}}''\right) = -\int_{\overline{\mathbf{r}}'}^{\overline{\mathbf{r}}''} \sum_{k=1}^{M} x_k\left(\mathbf{r}, y\right) dr_k. \tag{26}$$

This line integral, in general, is path dependent and does not have any economic meaning *per se*. Therefore, it cannot be a valid measure of welfare change. A remarkable case where the line integral is path independent and is a valid measure of welfare change is the one of quasi-linear preferences:

$$\begin{aligned} V(\mathbf{r}, y) &= a(\mathbf{r}) + y/r_0; \\ \frac{\partial x_k(\mathbf{r}, y)}{\partial y} &= 0, \quad k = 1...M; \\ x_0(\mathbf{r}, y) &= y - \sum_{k=1}^{M} p_k x_k(\mathbf{r}). \end{aligned} \tag{27}$$

Eq. (21) say that there is no income effect on the demand of all goods but the numéraire. In this case, the marginal utility of income is independent of income and of prices of all goods except the numéraire (the marginal utility of income cannot be constant for all values of income and prices):

$$\frac{\partial V(\mathbf{r}, y)}{\partial y} = \frac{1}{r_0}.$$

If we consider the change in prices $\overline{\mathbf{r}}' \to \overline{\mathbf{r}}''$, with quasi-linear preferences the Hotelling surplus line integral is path independent and is a valid measure of welfare change because it is proportional to the variation of the indirect utility:

$$\triangle S(\overline{\mathbf{r}}', \overline{\mathbf{r}}'') = -\int_{\overline{\mathbf{r}}'}^{\overline{\mathbf{r}}''} \sum_{k=1}^{M} x_k(\mathbf{r}, y)\, dr_k = r_0 \left[V(r_0, \overline{\mathbf{r}}'', y) - V(r_0, \overline{\mathbf{r}}', y)\right].$$

As shown by Williams (1976), the rule-of-a-half is the approximation of $\triangle S(\overline{\mathbf{r}}', \overline{\mathbf{r}}'')$ that is obtained from the Hotelling surplus line integral (26) if the path of integration is the segment between the points $\overline{\mathbf{r}}'$ and $\overline{\mathbf{r}}''$ and the demand functions are linearized:

$$\triangle S(\overline{\mathbf{r}}', \overline{\mathbf{r}}'') \cong \frac{1}{2} \sum_{k=1}^{M} \left[x_k(r_0, \overline{\mathbf{r}}', y) + x_k(r_0, \overline{\mathbf{r}}'', y)\right](r_k' - r_k'').$$

We define compensating variation CV the expenditure to be subtracted from the consumer's income in the state with the change to bring her to the state without the change:

$$V(r_0, \overline{\mathbf{r}}', y) = V(r_0, \overline{\mathbf{r}}'', y - CV).$$

We define equivalent variation EV the expenditure to be added to the consumer's income in the state without the change to bring her to the state with the change:

$$V(r_0, \overline{\mathbf{r}}', y + EV) = V(r_0, \overline{\mathbf{r}}'', y).$$

The compensating and equivalent variations do not suffer from the limitations of consumer's surplus: they are always valid measures of welfare change by definition. If preferences are quasi-linear we have:

$$CV = EV = \Delta S\left(\bar{\mathbf{r}}', \bar{\mathbf{r}}''\right).$$

We define the expenditure function of prices and utility:

$$e\left(\mathbf{r}, U\right) = \min\left\{\sum_{k=0}^{M} r_k x_k : U\left(\mathbf{x}\right) \geq U\right\},$$

and the compensated or Hicksian demand functions of prices and utility:

$$\mathbf{x}^c\left(\mathbf{r}, U\right) = \arg_{\mathbf{x}} \min\left\{\sum_{k=0}^{M} r_k x_k : U\left(\mathbf{x}\right) \geq U\right\}.$$

The compensating variation can also be expressed as the difference between the minimum expenditure needed to reach utility U' in the state without change with unchanged prices and the minimum expenditure needed to reach utility U' in the state without change with changed prices:

$$CV = e\left(r_0, \bar{\mathbf{r}}', U'\right) - e\left(r_0, \bar{\mathbf{r}}'', U'\right) = y - e\left(r_0, \bar{\mathbf{r}}'', U'\right).$$

Up to this point we have considered the individual consumer. We move now to the problem of aggregation.

Consider a population of consumers $n = 1, ..., N$ with heterogenous income y_n, $n = 1, ..., N$. Aggregate demand can be written as function of aggregate income $Y = \sum_{n=1}^{N} y_n$ if and only if the individual preferences have the Gorman polar form with the coefficient on income y_n the same for every consumer n, i.e. if and only if the indirect utility of the individual consumer has the form:

$$V_n = a_n\left(\mathbf{r}\right) + b\left(\mathbf{r}\right) y_n, \quad n = 1...N.$$

where $a_n\left(\mathbf{r}\right)$, $n = 1, ..., N$, and $b\left(\mathbf{r}\right)$ denote functions of $\mathbf{r}$.

It is said that a positive representative consumer exists. Quasi-linear preferences satisfy the Gorman polar form.

When the preferences of the individual consumers have the Gorman polar form with common coefficient $b\left(\mathbf{r}\right)$ and a utilitarian social welfare function is adopted, i.e. the sum of the indirect utilities of the individual consumers, then aggregate demand can be used to make welfare judgements by means of the techniques used for the individual consumer. It is said that a normative representative consumer exists. The aggregate indirect utility is in this case simply given by:

$$V\left(\mathbf{r}, Y\right) = \sum_{n=1}^{N} a_n\left(\mathbf{r}\right) + b\left(\mathbf{r}\right) Y.$$

Appendix B. Theoretical justification of econometric specifications

Different functional forms of the systematic utilities are used in applied work. One may wish to find a theoretical justification of the form that is used. Following the microeconomic foundation of discrete choice random utility models proposed by McFadden (1981), and further developed in Jara-Díaz and Farah (1988) and Jara-Díaz and Videla (1989), the systematic utility of each alternative is a conditional indirect utility.

The formulation of the model of consumer's behaviour is as follows. Assume the notation: U is the direct utility, $\mathbf{x} = (x_1, ..., x_M)$ is the row vector of the quantities of divisible goods, $\mathbf{r} = (r_1, ..., r_M)$ is the row vector of the prices of divisible goods, p_j is the price of the discrete alternative j, $\mathbf{q}_j$ is a vector of qualitative attributes of j, and y is income. Preliminarily, we notice that all prices and income are deflated by the price of a numéraire.

In the first step, we have the conditional on the choice of alternative j maximization problem:

$$\max_{\mathbf{x}} U(\mathbf{x}, j, \mathbf{q}_j), \tag{28}$$
$$\text{s.t.} \sum_{k=1}^{M} r_k x_k + p_j = y.$$

Assume now that the direct utility is additively separable in $(\mathbf{x}, j)$ and $\mathbf{q}_j$. The assumption implies $\partial^2 U / (\partial x_k \partial q_{js}) = 0, \ \forall k, s$. Thus we can re-write Eq. (22) as:

$$\max_{\mathbf{x}} U(\mathbf{x}, j, \mathbf{q}_j) = \max_{\mathbf{x}} U_1(\mathbf{x}, j) + U_2(\mathbf{q}_j). \tag{29}$$

Given the conditional demand functions:

$$\mathbf{x} = \mathbf{x}(\mathbf{r}, p_j, y),$$

by substitution into Eq. (29), we get the conditional indirect utility function:

$$V_j = V_{1j}(\mathbf{r}, p_j, y) + U_2(\mathbf{q}_j), \quad j = 1...J,$$

where J denotes the number of discrete alternatives.

To qualify as an indirect utility, the component $V_{1j}(\mathbf{r}, p_j, y)$ needs to be continuous in prices and income, non increasing in prices, non decreasing in income, homogenous of degree zero in prices and income, quasi-convex in prices and income.

Assume now that V_{1j} is additively separable in $\mathbf{r}$ and (p_j, y), with the component in $\mathbf{r}$ independent of the discrete alternative. We get:

$$V_j = a(\mathbf{r}) + w_j(p_j, y) + \overline{v}_j, \quad j = 1...J.$$

where $a(\mathbf{r})$ denotes a function of $\mathbf{r}$, $w_j(p_j, y)$ denotes a function of p_j and y, $\overline{v}_j = U_2(\mathbf{q}_j)$.

In the second step, we have the unconditional problem that seeks the alternative that maximizes the portion of V_j given by $v_j = w_j(p_j, y) + \overline{v}_j$. Notice that, by conditioning on the chosen alternative, $w_j(p_j, y)$ needs to satisfy Roy's identity:

$$-\frac{\partial w_j / \partial p_j}{\partial w_j / \partial y} = 1, \quad j = 1...J. \tag{30}$$

Roy's identity poses limitations on the functional form of w_j. Indeed, Eq. (30) can be written as:

$$\frac{\partial w_j}{\partial p_j} + \frac{\partial w_j}{\partial y} = 0, \quad j = 1...J. \tag{31}$$

Eq. (31) is the partial differential equation known in physics as advection or transport equation (when a substance is carried along a flow). A function $w_j(p_j, y)$ is a solution of Eq. (31) if and only if it satisfies (LeVeque, 2007; p. 202):

$$w_j = w_j\left[\pm\left(y - p_j\right)\right], \quad j = 1...J.$$

This means that we may admit only functional forms of w_j in plus/minus residual income $y - p_j$, a condition that may be violated by econometric specifications. This circumstance has led Viton (1985) to investigate the consistency between the microeconomic foundation and the econometric specification of random utility models, and to argue that, to restore consistency, income variables appearing in the specification may need to be re-interpreted, as an example as proxy for tastes rather than as true expenditure entering the budget constraint.

Acknowledgments

The authors would like to thank Lucas Javaudin for processing the Rhône data, which were then used for the estimation of the logit model. This study was supported by the NSFC-JPI UE joint research project 740 "MAAT" (project no. 18356856) and by the ANR projects ANR-11-LBX-0023-01 and ANR-16-IDEX-0008.

References

Anderson, S.P., de Palma, A., & Thisse, J.-F. 1988. A representative consumer theory of the logit model. *International Economic Review*, **29**(3), 461–466.

Anderson, S.P., de Palma, A., & Thisse, J.-F. 1992. *Discrete Choice Theory of Product Differentiation.* Cambridge MA: The MIT Press.

Ben-Akiva, M., & Lerman, S.R. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge MA: The MIT Press.

Bhattacharya, D. 2015. Nonparametric welfare analysis for discrete choice. *Econometrica*, **83**(2), 617–649.

Bhattacharya, D. 2018. Empirical welfare analysis for discrete choice: some general results. *Quantitative Economics*, **9**(2), 571–615.

Billingsley, P. 1995. *Probability and Measure*. Hoboken NJ: Wiley.

Chen, C., Tsaur, T., & Rhai, T. 1982. The Gini coefficient and negative income. *Oxford Economic Papers*, **34**(3), 473–478.

Cherchi, E., Polak, J., & Hyman, G. 2004. The impact of income, tastes and substitution effects on the assessment of user benefits using discrete choice models. *In: Proceedings of the European Transport Conference*.

Dagsvik, J.K., & Karlström, A. 2005. Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income. *Review of Economic Studies*, **72**(250), 57–76.

De Battisti, F., Porro, F., & Vernizzi, A. 2019. The Gini coefficient and the case of negative values. *Electronic Journal of Applied Statistical Analysis*, **12**(1), 85–107.

de Palma, A., & Kilani, K. 2005. Switching in the logit. *Economics Letters*, **88**(2), 196–202.

de Palma, A., & Kilani, K. 2011. Transition choice probabilities and welfare analysis in additive random utility models. *Economic Theory*, **46**(3), 427–454.

de Palma, A., Kilani, K, & Delle Site, P. 2021. *Welfare Inequality for Discrete Choices*. CY Cergy Paris Université, Mimeo.

Delle Site, P. 2008. On the geometry of the consumer's surplus line integral. *Economics Bulletin*, **4**(6), 1–7.

Delle Site, P. 2013. Integration of choice probabilities in logit. *Economics Letters*, **120**(1), 57–60.

Delle Site, P. 2021. A note on the large sample properties of the welfare change estimator in linear-in-income logit. *Journal of Choice Modelling*, **38**.

Delle Site, P., & Salucci, M.V. 2013. Transition choice probabilities and welfare analysis in random utility models with imperfect before-after correlation. *Transportation Research Part B: Methodological*, **58**, 215–242.

Delle Site, P., & Salucci, M.V. 2015. Transition choice probabilities in logit. *Economics Letters*, **126**, 135–139.

Galbis, A., & Maestre, M. 2012. *Vector Analysis versus Vector Calculus*. New York: Springer-Verlag.

Gastwirth, J.L. 1972. The estimation of the Lorenz curve and Gini index. *The Review of Economics and Statistics*, **54**(3), 306–316.

Herriges, J.A., & Kling, C.L. 1999. Nonlinear income effects in random utility models. *The Review of Economics and Statistics*, **81**(1), 62–72.

Hotelling, H. 1938. The general welfare in relation to problems of taxation and of railway and utility rates. *Econometrica*, **6**(3), 263–249.

Jara-Díaz, S. 1990. Consumer's surplus and the value of travel time savings. *Transportation Research Part B: Methodological*, **24**(1), 73–77.

Jara-Díaz, S. 2007. *Transport Economic Theory*. Amsterdam: Elsevier.

Jara-Díaz, S., & Farah, M. 1988. Valuation of users' benefits in transport systems. *Transport Reviews*, **8**(3), 197–218.

Jara-Díaz, S., & Videla, J. 1989. Detection of income effect in mode choice: theory and application. *Transportation Research Part B: Methodological*, **23**(6), 393–400.

Javaudin, L., Araldo, A., & de Palma, A. 2021. *Large-Scale Allocation of Personalized Incentives*. THEMA Working Papers 2021-08. THEMA (THéorie Economique, Modélisation et Applications), Université de Cergy-Pontoise.

Karlström, A. 2014. Appraisal. *Pages 601–626 of:* Hess, S., & Daly, A. (eds), *Handbook of Choice Modelling*. Cheltenham, UK: Edward Elgar.

Karlström, A., & Morey, E.R. 2004. *Calculating the exact compensating variation in logit and nested logit models with income effects: theory, intuition, implementation and application*. Paper presented at the Meeting of the American Economic Association, New Orleans, January 2001.

Kendall, M.G., & Stuart, A. 1958. *The Advanced Theory of Statistics: Distribution Theory*. London: Griffin.

Kotz, S., & Nadarajah, S. 2000. *Extreme Value Distributions. Theory and Applications*. London, UK: Imperial College Press.

LeVeque, R.J. 2007. *Finite Difference Methods for Ordinary and Partial Differential Equations - Steady-State and Time-Dependent Problems*. Philadelphia PA: SIAM, Society for Industrial and Applied Mathematics.

Ma, S., Kockelman, K.M., & Fagnant, D.J. 2015. Welfare analysis using logsum differences versus rule of half. Series of case studies. *Transportation Research Record: Journal of the Transportation Research Board*, **2530**, 73–83.

Mas-Colell, A., Whinston, M.D., & Green, J.R. 1995. *Microeconomic Theory*. New York: Oxford University Press.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Pages 105–142 of:* Zarembka, P. (ed), *Frontiers in Econometrics*. New York: Academic Press.

McFadden, D. 1978. Modelling the choice of residential location. *Pages 75–96 of:* Karlqvist, A., Lundqvist, L., Snickars, F., & Weibull, J. (eds), *Spatial Interaction Theory and Planning Models*. Amsterdam: North Holland.

McFadden, D. 1981. Econometric models of probabilistic choice. *Pages 198–272 of:* Manski, C.F., & McFadden, D. (eds), *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge MA: MIT Press.

McFadden, D. 1999. Computing willingness-to-pay in random utility models. *Pages 253–274 of:* Melvin, J.R., Moore, J.C., & Riezman, R. (eds), *Trade, Theory and Econometrics: Essays in Honor of John S. Chipman*. London: Routledge.

Mittelhammer, R.C. 2013. *Mathematical Statistics for Economics and Business*. New York: Springer.

Schutz, R.R. 1951. On the measurement of income inequality. *The American Economic Review*, **41**(1), 107–122.

Small, K.A., & Rosen, H.S. 1981. Applied welfare economics with discrete choice models. *Econometrica*, **49**(1), 105–130.

Takayama, A. 1994. *Analytical Methods in Economics*. Ann Arbor: University of Michigan Press.

Train, K. 2009. *Discrete Choice Methods with Simulation*. Cambridge UK: Cambridge University Press.

Tressider, J.O., Meyers, D.A., Burrell, J.E., & Powell, T.J. 1968. The London Transportation Study: methods and techniques. *Proceedings of the Institution of Civil Engineers*, **39**, 433–464.

Viton, P.A. 1985. On the interpretation of income variables in discrete choice models. *Economics Letters*, **17**(3), 203–206.

Wardman, M., Chintakayala, P., de Jong, G., & Ferrer, D. 2012. *European wide meta-analysis of values of travel time*. ITS University of Leeds. Paper prepared for EIB.

Williams, H.C.W.L. 1976. Travel demand models, duality relations and user benefits analysis. *Journal of Regional Science*, **16**(2), 147–166.

Williams, H.C.W.L. 1977. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A: Economy and Space*, **9**(3), 285–344.

Yang, H., & Huang, H.-J. 2005. *Mathematical and Economic Theory of Road Pricing*. Bingley, UK: Emerald.

Zhao, Y., Kockelman, K.M., & Karlström, A. 2012. Welfare calculations in discrete choice settings: an exploratory analysis of error term correlation with finite populations. *Transport Policy*, **19**(1), 76–84.